

### REMARKS

This application has been reviewed in light of the Office Action dated July 6, 2007. Claims 1-36 are pending in the application. By the present amendment, claims 21, 22, 25 and 32 have been amended. No new matter has been added. The Examiner's reconsideration of the rejection in view of the amendment and the following remarks is respectfully requested.

By the Office Action, the Examiner objected to claims 21, 22, and 25. While the applicant disagrees that any amendment need be made and that the claims are proper as presented, claims 21, 22 and 25 have been amended in a way believed to address the Examiner's concerns. Reconsideration is respectfully requested.

By the Office Action, the Examiner rejected claims 16 and 25 under 35 U.S.C. §112, second paragraph, as being indefinite. Claim 25 has been amended in a way believed to overcome the rejection. Reconsideration is respectfully requested.

By the Office Action, claims 1, 3, 5, 7, 15-16 and 31 stand rejected under 35 U.S.C. §102(e) as being anticipated by U.S. Patent Publication No. 2005/0071169 to Steinbiss (hereinafter Steinbiss).

Steinbiss is directed to a voice command system employed for controlling appliances. Steinbiss in accordance with the command provides a delay period before the command takes action in accordance with a recognized command (see paragraphs [0039], [0041]). The delay after the voice command is employed to reduce the chance that the command is misconstrued. Steinbiss does not separate commands from acoustic data as contemplated by the present claims. In fact, Steinbiss provides a system that is limited to

recognizing commands only. Acoustic data (as opposed to commands) is not disclosed or suggested for any application or operation in Steinbiss.

Claim 1 of the present disclosure, includes, *inter alia*, a method for extracting commands and acoustic data in a same utterance, including decoding at least one word in acoustic data representing an acoustic signal that comprises a human utterance and determining acoustic word boundaries within the acoustic data; extracting at least one command in a decoded utterance; and identifying acoustic data segments in the utterance based on the acoustic word boundaries.

Steinbiss fails to disclose or suggest at least a method for extracting commands and acoustic data in a same utterance. Further, nowhere in Steinbiss is identifying acoustic data segments in the utterance based on the acoustic word boundaries disclosed or suggested. Instead, Steinbiss recognizes commands after a delay period. Acoustic data segments are not identified nor are they searched for in an utterance.

In accordance with the present claims, commands and acoustic data are extracted from a same utterance. Acoustic data segments are also identified in addition to commands in the utterance based on the acoustic word boundaries. The Examiner states that FIG. 1 shows acoustic data segments t1 and tr. This region includes the bounds for an entire command utterance S. No extracting of commands and acoustic data in a same utterance is performed. This utterance S is recognized as a command only and is not evaluated for both data segments and commands. It is therefore respectfully submitted that Steinbiss fails to teach or suggest the present invention as claimed in claim 1.

Claim 16 recites, *inter alia*, a method for recognizing at least one command and at least one segment of acoustic voice data in a same utterance including decoding at least one word in voice data representing the acoustic signal that comprises a human utterance and determining the acoustic word boundaries within the voice data; extracting at least one command from the utterance; and associating segments in the voice data based on the acoustic word boundaries with labels.

Steinbiss does not disclose or suggest recognizing at least one command and at least one segment of acoustic voice data in a same utterance, as set forth in claim 16. While a command may be recognized in the utterance, acoustic voice data segments are not recognized or even searched for in Steinbiss. In addition, Steinbiss also fails to disclose or suggest anything about associating segments in the voice data based on the acoustic word boundaries with labels.

In accordance with claim 16, labels are employed to associate voice data segments. The labels are applied to the data to mark boundaries to indicate, e.g., whether the acoustic information should be saved as acoustic data (acoustic signals themselves) or decoded into text. Labels (e.g., 1000, 2000, 3000, etc) as described in detail in the specification (see page 26 for example) are employed to identify portions of the utterance which do not have to be decoded or have other meanings. Steinbiss does not disclose or suggest such labels, nor are the functions or application of such labels disclosed or suggested. Instead, Steinbiss provides an utterance that is recognized as a command. The beginning and end of the utterance are determined ( $t_1$  and  $t_2$ ), but no labels are employed in the manner described by claim 16.

Therefore, it is respectfully submitted that Steinbiss fails to disclose or suggest at

least a method for recognizing at least one command and at least one segment of acoustic voice data in a same utterance ... and associating segments in the voice data based on the acoustic word boundaries with labels. Steinbiss fails to teach or suggest that present invention as claimed in claim 16.

Accordingly, claims 1 and 16, and claims dependent therefrom (2-15 and 17-31) are believed to be in condition for allowance for at least the stated reasons.

By the Office Action, claims 2, 4, 6, 14, 18-20, 23 and 30 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Steinbiss in view of U.S. Patent No. 6,839,670 to Stammler et al. (hereinafter Stammler), claims 8-13, 17 and 24-29 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Steinbiss in view of U.S. Patent No. 6,434,529 to Walker et al. (hereinafter Walker), and claims 21 and 22 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Steinbiss in view of U.S. Patent No. 6,434,520 to Kanevsky et al. (hereinafter Kanevsky).

The Applicant disagrees with these rejections for at least the above stated reasons. Stammler, Walker and/or Kanevsky, taken alone or in any combination fail to cure the deficiencies of Steinbiss, as stated. It is further noted that Kanevsky is commonly owned with the present disclosure by International Business Machines Corporation. Reconsideration of the rejection is earnestly solicited.

By the Office Action, claims 32-36 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Walker in view of U.S. Patent Publication No. 2002/0111803 to Romero (hereinafter Romero).

Walker is directed to a system that references object instances to invoke a

method on the objects from within a speech recognition grammar. Character strings are mapped to object instances, e.g., in a table. When a reference to the table is made, a method or program is invoked. This eliminates the need for creating relationship between objects in logic. The motivation in Walker is to reduce costs of post-processing of feature/value tables in programs. (see col. 4, lines 8-62). The Examiner stated that Walker does not teach/mention:

a speech recognition engine configured to match portions of the utterances to acoustic models and language models to recognize words and word boundaries in the utterance and labels commands in the utterance; at least one program that executes label-identified commands and processes remaining portions of the utterance.

Romero was cited to cure these deficiencies. It should be noted that Romero is commonly assigned with the present disclosure to International Business Machines Corporations. Should the Examiner believe that an obvious-type double patenting rejection is in order, the Applicant would consider a terminal disclaimer.

Romeo is directed to a natural language understanding system that permits a dialog between a human and a machine. The speech recognition in Romero is geared towards "concepts and data". However, Romero fails to disclose or suggest at least: at least one program that executes label-identified commands and processes remaining portions of the utterance including processing audio data parts separately from the commands using a different vocabulary, the vocabulary being selected in accordance with at least one command in the utterance.

Claim 32 recites, *inter alia*, a system for recognizing commands and voice data in a same utterance including ... at least one program that executes label-identified commands

and processes remaining portions of the utterance including processing audio data parts separately from the commands using a different vocabulary, the vocabulary being selected in accordance with at least one command in the utterance.

The cited combination fails to disclose or suggest a program that executes label-identified commands and processes remaining portions of the utterance including processing audio data parts separately from the commands using a different vocabulary, the vocabulary being selected in accordance with at least one command in the utterance. By recognizing commands and voice data in a same utterance, commands can be executed that employ the acoustic data provided in the utterance. Additional information from the command can be used to switch to a different vocabulary to more accurately process the acoustic or voice data.

In an example given in the specification beginning at page 10:

“The partial utterance “His number is” can be decoded into a vocabulary change command and the partial utterance “six six six, one two three four” is voice data to be decoded into numbers. Note that this differs significantly from asking the user for the number in that the words “His number is” and equivalent phrases would not and should not be included in a number-recognition vocabulary because doing so would degrade recognition accuracy.

In the example for entering names in a speech activated telephone dialer, making the interaction with the user simpler and more completely user-driven increases the convenience of the voice interface. A method for dividing utterances into commands and voice data may provide other benefits as well. For example, people do not always produce perfect information-containing utterances. Dislocations, inadvertent noises and social components may also be uttered. For example, a user may have said “His number is ... wait a second ... “six six six, one two three four” or “His number is six six six, cough, cough, one two three four” or “His number is six one five, no, six six six, one two three four.” Processing the “data” part of the utterance separately from the “command” part also permits reprocessing the voice data to extract the maximum amount of useful information without forcing retries.”

It is respectfully submitted that the cited combination fails to disclose or suggest at least: a system for recognizing commands and voice data in a same utterance including ... at


least one program that executes label-identified commands and processes remaining portions of the utterance including processing audio data parts separately from the commands using a different vocabulary, the vocabulary being selected in accordance with at least one command in the utterance. Claims 32-36 are believed to be in condition for allowance for at least the stated reasons. Reconsideration of the rejection is earnestly solicited.

In view of the foregoing amendments and remarks, it is respectfully submitted that all the claims now pending in the application are in condition for allowance. Early and favorable reconsideration of the case is respectfully requested.

It is believed that no additional fees or charges are currently due. However, in the event that any additional fees or charges are required at this time in connection with the application, they may be charged to applicant's IBM Deposit Account No. 50-0510.

Respectfully submitted,

Dated: 10/9/07

By:   
James J. Bitetto  
Registration No. 40,513

**Mailing Address:**

**KEUSEY, TUTUNJIAN & BITETTO, P.C.**  
20 Crossways Park North, Suite 210  
Woodbury, NY 11797  
Tel: (516) 496-3868  
Fax: (516) 496-3869